



Barnett, S. M. (2014) Optical Dirac equation. New Journal of Physics, 16 (9). 093008. ISSN 1367-2630

Copyright © 2014 IOP Publishing Ltd and Deutsche Physikalische Gesellschaft

<http://eprints.gla.ac.uk/96800>

Deposited on: 08 September 2014

Enlighten – Research publications by members of the University of Glasgow_
<http://eprints.gla.ac.uk>

Optical Dirac equation

Stephen M Barnett

School of Physics and Astronomy, University of Glasgow, Glasgow G12 8QQ, UK

E-mail: stephen.barnett@glasgow.ac.uk

Received 30 May 2014, revised 6 August 2014

Accepted for publication 11 August 2014

Published 5 September 2014

New Journal of Physics **16** (2014) 093008

doi:[10.1088/1367-2630/16/9/093008](https://doi.org/10.1088/1367-2630/16/9/093008)

Abstract

We write the charge-free Maxwell equations in a form analogous to that of the Dirac equation for a free electron. This allows us to apply to light some of the ideas developed for the relativistic theory of the electron. Valuable insight is gained, thereby, into the forms of the optical spin and orbital angular momenta.

Keywords: Dirac equation, mechanical properties of light, relativistic quantum theory, optical angular momentum

1. Historical introduction

Maxwell's equations are remarkable in that they have survived the introduction of quantum theory and of both the special and general theories of relativity. The form in which the equations are expressed has changed many times, however. It is interesting to note that in his original paper, Maxwell used no fewer than twenty symbols to represent his various fields [1]. It was Heaviside who, in the words of Sir Edmund Whittaker, 'clear(ed) away this accumulation of rubbish' and gave us Maxwell's equations in their now familiar form [2]. For the free electromagnetic field, that is in the absence of charges and currents, we write these as

$$\begin{aligned}\nabla \cdot \mathbf{E} &= 0, \\ \nabla \cdot \mathbf{B} &= 0, \\ \nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t},\end{aligned}$$



Content from this work may be used under the terms of the [Creative Commons Attribution 3.0 licence](https://creativecommons.org/licenses/by/3.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

$$\nabla \times \mathbf{B} = \frac{\partial \mathbf{E}}{\partial t}. \quad (1)$$

Here we have used the natural system of units in which ϵ_0 and μ_0 are both unity, as is the speed of light. The quantities \mathbf{E} and \mathbf{B} are the (real) electric field and the magnetic field (or perhaps, more properly, the magnetic flux density). It is common, also, to write these in a manifestly Lorentz-invariant form by use of the field tensor $F^{\mu\nu}$ or, indeed, in general covariant form to allow for the description of non-inertial frames of motion or the effects of a gravitational field [3, 4].

We shall be interested in the links between Maxwell's equations and those arising in quantum mechanics and especially in relativistic quantum theory. Perhaps the first step in this direction was published in the earliest days of the old quantum theory, although such a link was not the initial aim. Riemann and Silberstein introduced the complex vector field

$$\mathbf{F} = \frac{1}{\sqrt{2}}(\mathbf{E} + i\mathbf{B}), \quad (2)$$

the effect of which is to reduce Heaviside's four equations into just two [5–8]

$$\begin{aligned} \nabla \cdot \mathbf{F} &= 0, \\ \nabla \times \mathbf{F} &= i \frac{\partial \mathbf{F}}{\partial t}. \end{aligned} \quad (3)$$

The Riemann–Silberstein vector appears in some textbooks [9, 10] and has been applied to a variety of problems in classical electromagnetism including calculating the fields associated with a moving charge [11]. Of these two equations, the first may be viewed as a constraint on the forms of the allowed Riemann–Silberstein vectors with the second providing the dynamics. We can write the second equation in a matrix form [12–16]

$$i \frac{\partial}{\partial t} \mathcal{F} = -i\boldsymbol{\tau} \cdot \nabla \mathcal{F}, \quad (4)$$

where \mathcal{F} is the column vector

$$\mathcal{F} = \begin{pmatrix} F_x \\ F_y \\ F_z \end{pmatrix} \quad (5)$$

and $\boldsymbol{\tau}$ is a vector, the components of which are the familiar spin-1 matrices [17]

$$\begin{aligned} \tau_x &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \\ \tau_y &= \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix} . \\ \tau_z &= \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{aligned} \quad (6)$$

When written in the form (4) Maxwell's equations are reminiscent of the Schrödinger and Dirac equations. This is yet more apparent if we multiply by \hbar and introduce the momentum operator, $\mathbf{p} = -i\hbar \nabla$

$$i\hbar \frac{\partial}{\partial t} \mathcal{F} = \boldsymbol{\tau} \cdot \mathbf{p} \mathcal{F} \quad (7)$$

and identify $\boldsymbol{\tau} \cdot \mathbf{p}$ as our Hamiltonian. It is interesting to note the similarity between this equation and the two component theory of Weyl, appropriate for neutrinos, which we can obtain by replacing the three-component \mathcal{F} with a two-component field and the spin-1 τ -matrices with the familiar Pauli matrices appropriate for spin- $\frac{1}{2}$ particles [18].

Although it is not our main aim, no account of this topic would be complete without a discussion of the photon wave function. It is not surprising, given the above analysis, that many authors have sought a wave function for the photon. Perhaps the most natural approach is to adopt the Riemann–Silberstein vector, or equivalently \mathcal{F} , for this purpose [12, 13, 15, 16, 19–21]. One objection we might raise to this is that \mathcal{F} has the dimensions of the square root of an energy density rather than the square root of a (probability) density normally associated with a wave function. It has been suggested that we can solve this problem by dividing \mathcal{F} by $\sqrt{\rho}$ or $(-\hbar^2 \nabla^2)^{1/4}$ so that each frequency component of the field is divided by $\sqrt{\hbar\omega}$ [22–26]. The problem with this, however, is that the resulting wave functions exhibit highly nonlocal behaviour and a number of other undesirable features [16]. We can take the Riemann–Silberstein vector as our wave function but it is necessary to get the dimensions right either by dividing by the mean energy or by introducing the reciprocal of the Hamiltonian into the definition of our scalar product. These issues are discussed further in the excellent review article by Białynicki-Birula [16].

It is not our main aim to investigate quantum optical effects and we neither require nor seek a photon wave function. Instead, we seek to exploit the analogy between Maxwell's equations and the Dirac equation to learn something about, *classical* optics and, in particular, the mechanical properties of light. There exist a variety of Dirac-like formulations of Maxwell's equations, but the most appropriate for our purposes is essentially that proposed by Darwin [27, 28] for a six-component 'spinor'¹ ψ :

$$\psi = \frac{1}{\sqrt{2}} \begin{pmatrix} E_x \\ E_y \\ E_z \\ iB_x \\ iB_y \\ iB_z \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbf{E} \\ i\mathbf{B} \end{pmatrix}. \quad (8)$$

¹ We use the term 'spinor' here to emphasize the analogy with Dirac spinors for the electron. Of course the strict meaning of the term is intimately associated with transformations, most especially the rotation and Lorentz groups, [29–31]. We do not claim to associate such properties with our ψ and do not associate it with a true spinor in this sense, but retain the term so as to emphasize the analogy with the electron.

Our time-dependent Maxwell equations then acquire the form

$$i\frac{\partial}{\partial t}\psi = \begin{pmatrix} 0 & -i\boldsymbol{\tau} \cdot \boldsymbol{\nabla} \\ -i\boldsymbol{\tau} \cdot \boldsymbol{\nabla} & 0 \end{pmatrix}\psi, \quad (9)$$

where each element in the matrix represents a 3×3 matrix. It is this equation that we shall refer to as the optical Dirac equation. It is important to state this clearly as there are a number of equations that could also justifiably be given this title; in addition to those already mentioned earlier, for example, we should also mention in this regard the existence of Dirac-like formulations of Maxwell's equations based on four-component spinors [32, 33].

We should note that our spinor (8) is not unique. We shall find the physically relevant mechanical properties of the fields are expressed in terms of products of ψ and its conjugate ψ^\dagger and so we can multiply ψ by any phase factor and leave the physical properties unchanged. This is analogous, of course, to the undefined global phase of the wavefunction in quantum mechanics. There is also a more subtle symmetry between the electric and magnetic fields, due to Heaviside and Larmor [34–36]. This is the observation that the free-field Maxwell equations (1) are invariant under the transformation

$$\begin{aligned} \mathbf{E} &\rightarrow \mathbf{E} \cos \theta + \mathbf{B} \sin \theta, \\ \mathbf{B} &\rightarrow \mathbf{B} \cos \theta - \mathbf{E} \sin \theta, \end{aligned} \quad (10)$$

where θ is any angle. It follows that physically meaningful properties of the field should also respect this symmetry [37] and this idea has been useful in deriving forms for the spin and orbital angular momenta of the field [38, 39] and for its helicity [40–45]. We note that a variant of this symmetry applies also for gravitational waves and leads to locally conserved quantities analogous to those for electromagnetic waves [46]. We can enact the transformation (10) on our spinor by means of a unitary operator in the form

$$\psi \rightarrow \tilde{\psi} = \begin{pmatrix} \mathbf{I} \cos \theta & -i\mathbf{I} \sin \theta \\ -i\mathbf{I} \sin \theta & \mathbf{I} \cos \theta \end{pmatrix}\psi \quad (11)$$

and it follows that, *any* spinor of this form will give equivalent results. The choice of spinor is, therefore, a conventional one and we opt to work with ψ as given in (8).

In this paper we apply to our optical Dirac equation some of the methods that have been developed for Dirac's electron equation. We find, in particular, that the Foldy–Wouthysen transformation and the elimination of optical, *Zitterbewegung* lead, in a natural way, to familiar slowly-varying quantities.

The paper is intended to present the new material in its historical context and for this reason well-known material is blended with the less familiar and with material we believe to be original. It may help the more experienced and widely-read reader, however, to give an indication of where the novel material is to be found. The optical Foldy–Wouthysen transformation has been adapted from one for particles with spin 0 or spin 1 [47] but has not, to my knowledge, been applied to optics. The associated splitting of the fields into positive and negative frequency parts and the consequent derivation of the paraxial wave equation is, I think, new as is the approach to deriving slowly-varying mechanical properties. The spin and orbital angular momenta, helicity, zilches and infra-zilches have been given previously, but the novelty here is to show that they all arise naturally from the Dirac formalism.

It is pleasing to find, in particular, that many of the hard-won properties of optical angular momentum appear naturally from the optical Dirac equation.

2. The Dirac equation: a reminder

The Dirac equation was formulated to provide a quantum theory of the electron and a major motivation for this work was the existence of ‘duplexity’ phenomena, in which the number of observed energy eigenstates for an electron in an atom is twice that given by Schrödinger theory [48]. We note with interest, in light of our investigation of optical angular momentum, Dirac’s observation that ‘the resultant orbital angular momentum of an electron moving in an orbit in a central field of force is not a constant’. In its place, it is the total angular momentum, comprised of both an orbital and a spin part that is conserved.

The Dirac equation is now a standard part of the physics curriculum and a discussion of it can be found in any of a number of now standard texts [49–53]. We present, in this section, only a brief review of this material as a precursor to studying the corresponding features of our optical Dirac equation. Further details can be found by consulting the cited texts. The electron Dirac equation has the form

$$i\hbar \frac{\partial}{\partial t} \psi = \boldsymbol{\alpha} \cdot \mathbf{p} \psi + \beta m \psi, \quad (12)$$

where m is the mass of the electron and we continue to work in units in which the speed of light is unity. The state ψ is a four-component spinor and the operators α and β are the 4×4 matrices

$$\begin{aligned} \alpha &= \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ \boldsymbol{\sigma} & 0 \end{pmatrix}, \\ \beta &= \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}, \end{aligned} \quad (13)$$

where σ is a vector, the components of which are the familiar Pauli matrices, 0 is the 2×2 zero matrix and I is the unit matrix. The square of each of these matrices is the unit matrix and they mutually anticommute

$$\begin{aligned} \alpha_i \alpha_j + \alpha_j \alpha_i &= 2\delta_{ij}, \\ \beta \alpha_i + \alpha_i \beta &= 0. \end{aligned} \quad (14)$$

We can identify the positive quantity $\rho = \psi^\dagger \psi$ with the probability density for the electron position by normalizing it so that integration over all space gives unity. This density obeys an associated continuity equation, which may be derived directly from the Dirac equation

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} = 0, \quad (15)$$

where \mathbf{j} is identified as the probability current

$$\mathbf{j} = \psi^\dagger \boldsymbol{\alpha} \psi. \quad (16)$$

This leads us to identify α as the velocity operator for the electron. The eigenvalues of the three operators, α_x , α_y and α_z are all ± 1 , which each correspond to the velocity of light. We can resolve this apparent conflict with experience by noting that a wave-packet formed by superposing energy eigenstates moves with the expected group velocity given by the ratio of the

average momentum and the average energy. There is a remnant of the velocity eigenvalues in the rapid oscillations associated with the positive- and negative-energy components of the state. This, *Zitterbewegung* has a frequency in excess of $2m/\hbar$ or, equivalently, a spatial period comparable to the Compton wavelength of the electron.

2.1. Mechanical properties

The properties of an electron in relativistic quantum theory are represented by operators, just as in the more familiar non-relativistic theory. The position and momentum operators, \mathbf{x} and \mathbf{p} satisfy the canonical commutation relation

$$[x_j, p_k] = i\hbar\delta_{jk} \quad (17)$$

and the Dirac equation (12) leads directly to the Hamiltonian operator

$$H = \boldsymbol{\alpha} \cdot \mathbf{p} + \beta m. \quad (18)$$

This, together with the Heisenberg equation of motion, provides the evolution equations for the operators representing mechanical properties of interest. The position operator, for example, satisfies the equation

$$\dot{\mathbf{x}} = \frac{i}{\hbar}[H, \mathbf{x}] = \boldsymbol{\alpha}, \quad (19)$$

which we recognize as the velocity operator. The momentum operator commutes with the Hamiltonian and so, unlike the velocity operator, it is a constant of the motion

$$\dot{\mathbf{p}} = \frac{i}{\hbar}[H, \mathbf{p}] = 0. \quad (20)$$

This demonstrates the conservation of momentum for a free electron.

It is with the angular momentum that the surprises begin, as the orbital angular momentum for a free electron is not conserved. We can define an orbital angular momentum operator precisely as in non-relativistic quantum theory

$$\mathbf{L} = \mathbf{x} \times \mathbf{p}. \quad (21)$$

This operator does not commute with the Dirac Hamiltonian and so satisfies the non-trivial equation of motion

$$\dot{\mathbf{L}} = \frac{i}{\hbar}[H, \mathbf{L}] = \boldsymbol{\alpha} \times \mathbf{p}. \quad (22)$$

In contrast with the non-relativistic theory, the orbital angular momentum of a free electron, or of one confined by a rotationally symmetric potential, is not a constant of the motion. Dirac's resolution of this paradoxical situation was to add to the orbital angular momentum a spin angular momentum

$$\boldsymbol{\Sigma} = \frac{\hbar}{2} \begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix}. \quad (23)$$

Like the orbital angular momentum, this is not a constant of the motion, but satisfies the equation of motion

$$\dot{\boldsymbol{\Sigma}} = -\boldsymbol{\alpha} \times \mathbf{p}. \quad (24)$$

In relativistic quantum theory the spin and orbital angular momenta are naturally coupled together and it is only the total angular momentum, given by the sum of these, that is conserved. The total angular momentum operator

$$\mathbf{J} = \mathbf{L} + \mathbf{\Sigma} \quad (25)$$

is a constant of the motion. It is this quantity, rather than the orbital or spin angular momenta, that appears in the relativistic expressions for atomic spectra [50, 51, 54]. We should note that the separation implicit in (25) is not, itself, Lorentz covariant.

2.2. Foldy–Wouthuysen transformation

It is apparent that the velocity operator for a Dirac electron is very different from that in non-relativistic quantum theory. This feature, together with the non-conservation of spin and orbital angular momentum and a desire to understand the non-relativistic limit motivated the discovery of a unitary transformation that diagonalizes the Dirac Hamiltonian [55]. The required unitary operator is generated by an Hermitian and time-independent operator, S , and has the form [51, 55]

$$e^{iS} = \exp(\theta \beta \boldsymbol{\alpha} \cdot \mathbf{p}) = \cos(p\theta) + \frac{\beta \boldsymbol{\alpha} \cdot \mathbf{p}}{p} \sin(p\theta), \quad (26)$$

where $p = |\mathbf{p}|$ is the magnitude of the momentum and θ is a parameter to be determined. It is not difficult to find that the required value of θ is

$$\theta = \frac{1}{2p} \tan^{-1}\left(\frac{p}{m}\right). \quad (27)$$

The Foldy–Wouthuysen transformation produces a new state and Dirac equation of the form

$$\begin{aligned} \psi' &= e^{iS} \psi \\ i\hbar \frac{\partial}{\partial t} \psi' &= H' \psi', \end{aligned} \quad (28)$$

where H' is the transformed (diagonal) Hamiltonian

$$H' = e^{iS} H e^{-iS} = \beta \sqrt{m^2 + p^2}. \quad (29)$$

The operator β has eigenvalues ± 1 , with the positive eigenvalue corresponding to the upper two components of ψ and the negative to the lower two. Clearly, the transformation has succeeded in separating the positive energy states from those with negative energy. The eigenvalues of H' , moreover, are $\pm \sqrt{m^2 + p^2}$ in agreement with the relationship between energy, mass and momentum, $E^2 = m^2 + p^2$, familiar from relativistic kinematics.

Our transformed Dirac equation has the form

$$i\hbar \frac{\partial}{\partial t} \psi' = \beta \sqrt{m^2 + p^2} \psi'. \quad (30)$$

If we consider only low energies in the non-relativistic regime then this, on making a binomial expansion, reduces to

$$i\hbar \frac{\partial}{\partial t} \psi' \approx \left(m - \frac{\hbar^2}{2m} \nabla^2 \right) \beta \psi'. \quad (31)$$

If we further restrict our attention only to the two-component positive energy part of the state and shift our zero of energy, so as to remove the constant m term, then we recover the familiar Schrödinger equation for a free particle.

The operators representing the mechanical and other properties of the electron are also transformed by the Foldy–Wouthuysen transformation. In particular, the position operator \mathbf{x} is transformed to [55]

$$\begin{aligned}\mathbf{x}' &= e^{iS}\mathbf{x}e^{-iS} \\ &= \mathbf{x} - \frac{i\hbar\beta\boldsymbol{\alpha}}{2E_p} - \frac{i\mathbf{p}\hbar\beta\boldsymbol{\alpha} \cdot \mathbf{p} - 2p\boldsymbol{\Sigma} \times \mathbf{p}}{2E_p(E_p + m)p},\end{aligned}\quad (32)$$

where E_p is the operator $\sqrt{m^2 + p^2}$. The momentum operator commutes with S and so is unchanged by the transformation. The orbital and spin angular momenta are also changed by the Foldy–Wouthuysen transformation, with the orbital angular momentum operator becoming

$$\mathbf{L}' = \mathbf{x}' \times \mathbf{p}. \quad (33)$$

The total angular momentum commutes with S and so is unchanged by the transformation. We can readily use this fact to obtain the transformed spin angular momentum operator. It is important to realize that using the transformed Hamiltonian, states and observables is entirely equivalent to working with the untransformed quantities and the original Dirac equation.

2.3. Mean properties

In the non-relativistic limit the position and momentum operators are \mathbf{x} and \mathbf{p} , as they are in the Dirac theory. With this non-relativistic limit in mind, it is reasonable to ask what role these operators play in the transformed system. We note that in the transformed picture, the original position operator satisfies the equation of motion

$$\dot{\mathbf{x}} = \frac{i}{\hbar}[H', \mathbf{x}] = \frac{\mathbf{p}(\beta m + \boldsymbol{\alpha} \cdot \mathbf{p})}{m^2 + p^2}. \quad (34)$$

For a positive energy state this gives the velocity operator we might have expected: \mathbf{p}/E_p . For a negative energy state the velocity is the same apart from an overall minus sign. The, *Zitterbewegung* or rapid oscillations, associated with an instantaneous velocity equal to that of light, has disappeared. For this reason, Foldy and Wouthuysen refer to the operator \mathbf{x} in their transformed system as the, *mean-position* operator. Working with this mean-position operator, \mathbf{x} , and the transformed Dirac equation (30) gives the familiar Schrödinger evolution in the non-relativistic limit. The other original operators \mathbf{L} and $\boldsymbol{\Sigma}$ are similarly the, *mean-orbital* and *mean-spin* angular momenta. It is important to note that it is these mean quantities that correspond to the those familiar from non-relativistic quantum theory [55].

The mean quantities can also be used with the original Dirac equation by simply inverting the Foldy–Wouthuysen transformation. This means, in particular, that the mean position, orbital and spin angular momentum operators appropriate for use with the Dirac equation are, respectively,

$$\begin{aligned}
\bar{\mathbf{x}} &= e^{-iS} \mathbf{x} e^{iS}, \\
\bar{\mathbf{L}} &= e^{-iS} \mathbf{L} e^{iS}, \\
\bar{\mathbf{\Sigma}} &= e^{-iS} \mathbf{\Sigma} e^{iS},
\end{aligned} \tag{35}$$

echoing the quantities introduced by Pryce [56, 57]. The mean orbital and spin angular momenta, unlike their unaveraged counterparts, commute with the Hamiltonian and so are true constants of the motion. In the Foldy–Wouthuysen picture this means that

$$[H', \mathbf{L}] = 0 = [H', \mathbf{\Sigma}] \tag{36}$$

or, for the original Dirac Hamiltonian

$$[H, \bar{\mathbf{L}}] = 0 = [H, \bar{\mathbf{\Sigma}}]. \tag{37}$$

3. Properties of the optical Dirac equation

It is natural to ask if the Dirac equation can be applied to particles with spins other than the $\frac{1}{2}$ associated with the electron [58] and an early application was to the (bosonic) mesons [59]. Our interest, of course, is the representation of the electromagnetic waves in Dirac form and to this end we rewrite our optical Dirac equation (9) in the form

$$i\hbar \frac{\partial}{\partial t} \psi = \boldsymbol{\alpha} \cdot \mathbf{p} \psi, \tag{38}$$

where we associate the ‘momentum’ operator \mathbf{p} with $-i\hbar \nabla$ and $\boldsymbol{\alpha}$ is a vector the components of which are the 6×6 matrices

$$\boldsymbol{\alpha} = \begin{pmatrix} 0 & \boldsymbol{\tau} \\ \boldsymbol{\tau} & 0 \end{pmatrix}. \tag{39}$$

To these three matrices we add the 6×6 form of Dirac’s β -matrix

$$\beta = \begin{pmatrix} \mathbf{I} & 0 \\ 0 & -\mathbf{I} \end{pmatrix}. \tag{40}$$

We have deliberately written these matrices using the same notation as that for the electron Dirac equation, so as to emphasize the similarities between them. The differences are apparent in the fact that the matrices for the optical Dirac equation are 6×6 , rather than the 4×4 for the electron, and also that the commutation relations among the matrices are rather different [47, 59–65]. These are summarized in appendix A.

The optical Dirac equation leads to the same conservation law as for its electronic counterpart (15) but the relevant quantities are now the energy density $w = \psi^\dagger \psi = \frac{1}{2}(E^2 + B^2)$ and the Poynting vector $\mathbf{S} = \psi^\dagger \boldsymbol{\alpha} \psi = \mathbf{E} \times \mathbf{B}$

$$\begin{aligned}
\frac{\partial}{\partial t} (\psi^\dagger \psi) &= -(\nabla \psi^\dagger) \cdot \boldsymbol{\alpha} \psi - \psi^\dagger \boldsymbol{\alpha} \cdot (\nabla \psi) \\
\Rightarrow \frac{\partial}{\partial t} w + \nabla \cdot \mathbf{S} &= 0,
\end{aligned} \tag{41}$$

which we recognise as Poynting’s theorem. The difference in the dimensions of ψ for the electron and optical Dirac equations is manifested, here, in the change of meaning of the continuity equation from the local conservation of probability to the local conservation of energy.

3.1. Mechanical properties of light

The derivation of Poynting's theorem as the continuity equation for the optical Dirac equation led us to identify the density of energy and the flux of energy, or the momentum density, as

$$\begin{aligned} w &= \psi^\dagger \psi \\ \mathbf{S} &= \psi^\dagger \boldsymbol{\alpha} \psi. \end{aligned} \quad (42)$$

To these we can add the density of the optical angular momentum, which we write as

$$\mathbf{j} = \psi^\dagger (\mathbf{r} \times \boldsymbol{\alpha}) \psi = \mathbf{r} \times (\mathbf{E} \times \mathbf{B}). \quad (43)$$

We trust that no confusion will arise if we stick with established convention in which both the optical angular-momentum density and the electron probability current are represented by \mathbf{j} . Each of these mechanical quantities is globally conserved, of course, and we should confirm this fact using the properties of our optical Dirac equation. The global conservation of energy follows directly from the local conservation law (41)

$$\begin{aligned} \frac{\partial}{\partial t} \int d^3x \psi^\dagger \psi &= - \int d^3x \nabla \cdot (\psi^\dagger \boldsymbol{\alpha} \psi) \\ &= - \int \psi^\dagger \boldsymbol{\alpha} \psi \cdot d\mathbf{s} \\ &= 0, \end{aligned} \quad (44)$$

where we have used Gauss's theorem and the physical requirement that the fields tend to zero at infinity.

For the remaining properties we can proceed in the same way, using the Dirac equation for ψ and for ψ^\dagger . Alternatively, we can exploit further the analogy between the (classical) optical Dirac equation and the (quantum) Dirac equation for the electron, by introducing the 'Hamiltonian' operator

$$H = \boldsymbol{\alpha} \cdot \mathbf{p} = -i\hbar \boldsymbol{\alpha} \cdot \nabla \quad (45)$$

and use Heisenberg's equation of motion. Let us emphasize that this is an entirely equivalent procedure to working with the optical Dirac equation itself and that, despite the appearance of \hbar , the description is essentially classical. The optical momentum, for example is, expressed in terms of the operator $\boldsymbol{\alpha}$ and the momentum, *density* at position \mathbf{r} therefore corresponds to the operator $\boldsymbol{\alpha} \delta(\mathbf{r} - \mathbf{x})$, where \mathbf{x} is the position operator, as before [66–68]. Both approaches lead to the same result. For the electromagnetic momentum density we find

$$\begin{aligned} \frac{\partial}{\partial t} (\psi^\dagger \alpha_i \psi) &= - (\nabla_j \psi^\dagger) \alpha_j \alpha_i \psi - \psi^\dagger \alpha_i \alpha_j (\nabla_j \psi), \\ \Rightarrow \frac{\partial}{\partial t} S_i + \nabla_j T_{ji} &= 0, \end{aligned} \quad (46)$$

where we have introduced the summation convention in which a repeated index implies a summation over the three cartesian coordinates and T_{ji} is the familiar momentum flux density [36]

$$T_{ji} = \frac{1}{2} \delta_{ij} (E^2 + B^2) - E_i E_j - B_i B_j. \quad (47)$$

In deriving the momentum continuity equation (46) we use the commutation properties of the Kemmer matrices, given in appendix A and the transverse nature of \mathbf{E} and \mathbf{B} which mean that $\mathbf{p} \cdot \mathbf{E} = 0 = \mathbf{p} \cdot \mathbf{B}$. If we follow the same procedure for the operator $\mathbf{x} \times \boldsymbol{\alpha}$ then we find the familiar local conservation for the electromagnetic angular momentum

$$\begin{aligned} \frac{\partial}{\partial t} (\psi^\dagger (\mathbf{x} \times \boldsymbol{\alpha})_i \psi) &= -(\nabla_j \psi^\dagger) \alpha_j \psi^\dagger (\mathbf{x} \times \boldsymbol{\alpha})_i \psi - \psi^\dagger (\mathbf{x} \times \boldsymbol{\alpha})_i \alpha_j (\nabla_j \psi), \\ \Rightarrow \frac{\partial}{\partial t} j_i + \nabla_j M_{ji} &= 0, \end{aligned} \quad (48)$$

where \mathbf{j} is given by (43) and M_{ji} is the angular-momentum flux density [36, 69]

$$M_{ji} = \varepsilon_{ilk} x_l T_{jk}. \quad (49)$$

3.2. Optical Foldy–Wouthuysen transformation

The expressions obtained in the preceding subsection are exact (within the Maxwell theory of the free field) and contain both rapidly and slowly oscillating contributions. For a monochromatic field of angular frequency ω , for example, the densities and flux densities of the energy, momentum and angular momentum include constant terms and terms oscillating at frequency 2ω . In the majority of situations in optics, this frequency is too high to be observed in the interaction with matter and it is appropriate to remove the rapidly oscillating contributions so as to arrive at slowly-varying or mean properties. This process is known by different terms in different parts of optics: we may think of it as averaging over many optical cycles, the slowly-varying amplitude approximation or, within quantum optics, as making the rotating-wave approximation in the interaction with a detector. The situation is reminiscent and indeed precisely analogous to the, *Zitterbewegung* apparent in the physical quantities calculated from the Dirac equation for the electron. It is natural to construct the appropriate cycle-averaged optical properties by adopting the same procedure as for the electron: we apply a Foldy–Wouthuysen transformation to diagonalize the Hamiltonian for our optical Dirac equation and then seek suitable operators from which to obtain mean mechanical properties.

The optical Foldy–Wouthuysen transformation we seek is the natural analogue of that for the electron [47] with the required unitary operator having the form

$$e^{iS} = \exp(\theta \beta \boldsymbol{\alpha} \cdot \mathbf{p}), \quad (50)$$

where θ is a function, to be determined, of the, *operator* $p = |\mathbf{p}|$. We can expand the unitary operator (50) in a form analogous to that given for the electron transformation in (26), although the form is complicated, somewhat, by the algebra of the Kemmer, as opposed to Dirac matrices. We find the form

$$e^{iS} = \frac{\mathcal{P}^2}{p^2} + \left(1 - \frac{\mathcal{P}^2}{p^2}\right) \cos(p\theta) + \frac{\beta \boldsymbol{\alpha} \cdot \mathbf{p}}{p} \sin(p\theta), \quad (51)$$

where we define the operator

$$\mathcal{P}^2 = p^2 - (\boldsymbol{\alpha} \cdot \mathbf{p})^2, \quad (52)$$

the properties of which are given in appendix A.

The desired diagonalized Hamiltonian is obtained by setting $\theta = \pi/(4p)$, as shown in appendix B

$$\begin{aligned} H' &= e^{iS} H e^{-iS} \\ &= \beta \left(p - \frac{p^2}{p} \right). \end{aligned} \quad (53)$$

It follows that the eigenvalues of the Hamiltonian are $\pm p$, corresponding to the wavenumber or angular frequency $\omega = p/\hbar$, as $p^2\psi = 0$.

To complete the transformation we need to find the transformed spinor

$$\begin{aligned} \psi' &= e^{iS}\psi = \exp\left(\frac{\pi}{4p}\beta\boldsymbol{\alpha} \cdot \mathbf{p}\right)\psi \\ &= \frac{1}{\sqrt{2}}\left(1 + \frac{\beta\boldsymbol{\alpha} \cdot \mathbf{p}}{p}\right)\psi. \end{aligned} \quad (54)$$

We can write this in terms of our electric and magnetic fields if we introduce complex electric and magnetic fields corresponding to the positive, $\mathbf{E}^{(+)}$, $\mathbf{B}^{(+)}$, and negative, $\mathbf{E}^{(-)}$, $\mathbf{B}^{(-)}$, frequency parts of the field

$$\begin{aligned} \mathbf{E} &= \mathbf{E}^{(+)} + \mathbf{E}^{(-)} \\ \mathbf{B} &= \mathbf{B}^{(+)} + \mathbf{B}^{(-)}, \end{aligned} \quad (55)$$

where

$$\begin{aligned} \hbar \nabla \times \mathbf{E}^{(\pm)} &= \pm ip\mathbf{B}^{(\pm)} \\ \hbar \nabla \times \mathbf{B}^{(\pm)} &= \mp ip\mathbf{E}^{(\pm)}. \end{aligned} \quad (56)$$

These positive and negative frequency parts of the field are those that arise naturally in the theory of optical coherence and are associated, in the quantum theory of light, with the photon annihilation and creation operators respectively [70–72]. The transformed spinor naturally separates these components with positive frequency electric components in the top three entries and negative frequency components in the lower three

$$\psi' = \begin{pmatrix} \mathbf{E}^{(+)} \\ i\mathbf{B}^{(-)} \end{pmatrix}. \quad (57)$$

It is useful to note that the action of $\boldsymbol{\alpha} \times \mathbf{p}/p$ on our transformed spinor gives the minus complex conjugate

$$\frac{\boldsymbol{\alpha} \cdot \mathbf{p}}{p}\psi' = -\psi'^*, \quad (58)$$

and similary

$$\psi'^{\dagger} \frac{\boldsymbol{\alpha} \cdot \mathbf{p}}{p} = -\psi'^{* \dagger} = -\psi'^T. \quad (59)$$

We emphasize that the combination of this transformed spinor and the transformed Hamiltonian, H' , is exact and fully equivalent to our untransformed quantities.

We conclude this subsection by noting that the paraxial approximation, a mainstay of laser physics [73, 74], is readily derived from our transformed optical Dirac equation. It follows from (53) and the transverse nature of the fields that our transformed Dirac equation has the form

$$i\hbar \frac{\partial}{\partial t} \psi' = \beta p \psi'. \quad (60)$$

There is no mass term, as in the corresponding equation for the electron (30), but we can arrive at a similar Schrödinger-like equation if one component of the momentum, which we take to be the z -component, is far larger than the others, so that $p_z \approx p = E$. In this case we write

$$p = \sqrt{p_z^2 + p_x^2 + p_y^2} \approx p_z + \frac{p_x^2 + p_y^2}{2p}. \quad (61)$$

If we make the ansatz

$$\psi' = e^{ip(\beta z - t)/\hbar} \psi'' \quad (62)$$

and multiply the equation by β , then our transformed Dirac equation becomes

$$i \left(\frac{\partial}{\partial z} + \beta \frac{\partial}{\partial t} \right) \psi'' = -\frac{\hbar}{2p} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) \psi'', \quad (63)$$

so that cartesian components of the positive and negative frequency parts of both the electric and magnetic fields obey, in this approximation, a suitable paraxial wave equation [73, 74].

3.3. Mean optical properties

We can use our transformed spinor, ψ' , to arrive at mean mechanical properties by following the procedure Foldy and Wouthuysen applied to the electron Dirac equation [55]. We transform the operators corresponding to the mechanical properties of interest and drop terms that are of odd order in the α , as these introduce cross terms between the positive and negative frequency parts of the spinor. This serves to remove the rapidly oscillating contributions in the same way as the corresponding procedure for the electron removes the, *Zitterbewegung*.

Let us begin by considering the energy density. We have transformed the spinor and our form for the energy density must be unchanged if we also transform the operator the expectation value of which is the energy density. It is straightforward to see that the operator for the energy density at position \mathbf{r} must be $\delta(\mathbf{r} - \mathbf{x})$, where \mathbf{x} is the position operator, as

$$w(\mathbf{r}) = \int d^3x \psi^\dagger \delta(\mathbf{r} - \mathbf{x}) \psi. \quad (64)$$

Naturally, we obtain exactly the same result if we work with our transformed spinor ψ' and the transformed energy density operator

$$\begin{aligned} w(\mathbf{r}) &= \int d^3x \psi'^{\dagger} e^{iS} \delta(\mathbf{r} - \mathbf{x}) e^{-iS} \psi' \\ &= \frac{1}{2} \int d^3x \psi'^{\dagger} \left(1 + \frac{\beta \boldsymbol{\alpha} \cdot \mathbf{p}}{p} \right) \delta(\mathbf{r} - \mathbf{x}) \left(1 - \frac{\beta \boldsymbol{\alpha} \cdot \mathbf{p}}{p} \right) \psi', \end{aligned} \quad (65)$$

where we have used the fact that both ψ' and ψ'^{\dagger} contain only transverse fields so that $\mathcal{P}^2\psi' = 0$ and $\psi'^{\dagger}\mathcal{P}^2 = 0$. We obtain the mean or cycle-averaged energy density by keeping only those terms that are of even order in α

$$\begin{aligned}\bar{w}(\mathbf{r}) &= \frac{1}{2} \int d^3x \psi'^{\dagger} \delta(\mathbf{r} - \mathbf{x}) \psi' - \frac{1}{2} \int d^3x \psi'^{\dagger} \frac{\beta \boldsymbol{\alpha} \cdot \mathbf{p}}{p} \delta(\mathbf{r} - \mathbf{x}) \frac{\beta \boldsymbol{\alpha} \cdot \mathbf{p}}{p} \psi' \\ &= \frac{1}{2} \left(\psi'^{\dagger}(\mathbf{r}) \psi'(\mathbf{r}) + \psi'^{* \dagger}(\mathbf{r}) \psi'^*(\mathbf{r}) \right) \\ &= \psi'^{\dagger}(\mathbf{r}) \psi'(\mathbf{r}) \\ &= \mathbf{E}^{(-)} \cdot \mathbf{E}^{(+)} + \mathbf{B}^{(-)} \cdot \mathbf{B}^{(+)},\end{aligned}\tag{66}$$

which is the required average. In reaching this result we have used the anti-commutation relation (A.1) and the complex-conjugation properties (58) and (59). This simple expression for the cycle-averaged energy density illustrates the benefit of working with the transformed spinor.

We can continue in the same way, by keeping terms of only even order in α to find the averaged forms of our other mechanical properties. For the momentum density we find

$$\begin{aligned}\bar{\mathbf{S}} &= \frac{1}{2} \int d^3x \psi'^{\dagger} \left[\frac{\beta \boldsymbol{\alpha} \cdot \mathbf{p}}{p}, \boldsymbol{\alpha} \delta(\mathbf{r} - \mathbf{x}) \right] \psi' \\ &= -\frac{1}{2} \int d^3x \left[\psi'^{\dagger} \beta \boldsymbol{\alpha} \delta(\mathbf{r} - \mathbf{x}) \psi' + \psi'^{\dagger} \delta(\mathbf{r} - \mathbf{x}) \boldsymbol{\alpha} \beta \psi'^* \right] \\ &= \mathbf{E}^{(+)} \times \mathbf{B}^{(-)} + \mathbf{E}^{(-)} \times \mathbf{B}^{(+)},\end{aligned}\tag{67}$$

which is the required cycle-averaged form of Poynting's vector. In reaching this expression, we have made use of the readily proved identity

$$\mathbf{a} \cdot \boldsymbol{\alpha} \psi' = \begin{pmatrix} -\mathbf{a} \times \mathbf{B}^{(-)} \\ i\mathbf{a} \times \mathbf{E}^{(+)} \end{pmatrix}.\tag{68}$$

The averaged momentum density (66), although evidently correct, is not in its most useful form. We can obtain a form closer to that used in optics [75] by evaluating the operator product $(\boldsymbol{\alpha} \times \mathbf{p})\boldsymbol{\alpha}$ using (A.9). After a little work we find

$$\begin{aligned}\bar{\mathbf{S}} &= \frac{1}{2} \left[E_i^{(-)} \frac{\overleftarrow{\mathbf{p}}}{p} E_i^{(+)} - B_i^{(+)} \frac{\overrightarrow{\mathbf{p}}}{p} B_i^{(-)} + E_i^{(-)} \frac{\overrightarrow{\mathbf{p}}}{p} E_i^{(+)} - B_i^{(+)} \frac{\overleftarrow{\mathbf{p}}}{p} B_i^{(-)} \right. \\ &\quad - \mathbf{E}^{(-)} \left(\frac{\overleftarrow{\mathbf{p}}}{p} \cdot \mathbf{E}^{(+)} \right) + \mathbf{B}^{(+)} \left(\frac{\overleftarrow{\mathbf{p}}}{p} \cdot \mathbf{B}^{(-)} \right) \\ &\quad \left. - \left(\mathbf{E}^{(-)} \cdot \frac{\overrightarrow{\mathbf{p}}}{p} \right) \mathbf{E}^{(+)} + \left(\mathbf{B}^{(+)} \cdot \frac{\overrightarrow{\mathbf{p}}}{p} \right) \mathbf{B}^{(-)} \right].\end{aligned}\tag{69}$$

We recall that $\vec{\mathbf{p}} = -i\hbar \nabla$ and $\overleftarrow{\mathbf{p}} = i\hbar \nabla$ and so write this in the simpler form

$$\begin{aligned} \bar{\mathbf{S}} = & \text{Im} \left[E_i^{(-)} \left(\frac{\nabla}{\omega} E_i^{(+)} \right) + B_i^{(-)} \left(\frac{\nabla}{\omega} B_i^{(+)} \right) \right] \\ & + \nabla_i \text{Im} \left[E_i^{(+)} \left(\frac{\mathbf{E}^{(-)}}{\omega} \right) + B_i^{(+)} \left(\frac{\mathbf{B}^{(-)}}{\omega} \right) \right]. \end{aligned} \quad (70)$$

We emphasize that here ω is an, *operator* and that no restriction to monochromatic or near-monochromatic fields is necessary. The action of the operator $1/\omega$ simply divides each temporal Fourier component of the field by its associated angular frequency. The second line in (70) is a divergence and therefore does not contribute to the total momentum density. It should be noted that this operator introduces some spatial averring in that, for example, $\mathbf{B}^{(-)}(\mathbf{r})/\omega$ is not simply related to \mathbf{B} at position \mathbf{r} alone. The density of a mechanical property is only defined up to such a divergence and so we are at liberty to neglect this term. If this statement needs justification, we recall that conserved quantities, like momentum, may be obtained from symmetries by means of Noether's theorem and that the local resulting local conservation law only, *defines* the density up to a divergence [76–79]². We note, however, that retaining such divergence terms allows us to express the momentum density in a variety of suggestive forms [80], although, because of the non-uniqueness described above, it is difficult to assign to these any particular physical significance. If we drop the divergence term then we are left with the simple form

$$\begin{aligned} \bar{\mathbf{S}} = & \text{Im} \left[E_i^{(-)} \frac{\nabla}{\omega} E_i^{(+)} + B_i^{(-)} \frac{\nabla}{\omega} B_i^{(+)} \right] \\ = & \text{Im} \left(\psi'^{\dagger} \frac{\nabla}{\omega} \psi' \right), \end{aligned} \quad (71)$$

which is reminiscent of the momentum density in non-relativistic quantum theory [81], although the presence of the frequency operator, ω is an important difference.

4. Optical angular momentum

Our task, in this section, is to determine how the optical angular momentum, the orbital and spin components and the helicity arise from our optical Dirac equation. There are three reasons for undertaking this task. Firstly, of course, our discussion of the mechanical effects of light would not be complete without including angular momentum. Secondly angular momentum played an important role in the development of Dirac's theory of the electron, in particular in determining the necessity of the existence of electron spin and its intimate relationship to the more familiar orbital angular momentum [48]. Finally optical angular momentum is currently a vibrant and

² It is beholden on me to alert the reader to two significant errors in the paper by Bliokh *et al* [79]. Firstly it is stated in their equation (3.36) that the, *total* spin and orbital angular momenta obtained from the conventional and dual-symmetric Lagrangians are different. They are not, as shown explicitly in [38]. Secondly, in their, *note added*, the authors, referring to the conventional and dual-symmetric Lagrangians, state that the choice makes a difference and has important physical consequences. It does not; both lead to the same Maxwell equations and are, therefore, *physically indistinguishable*. They have the same symmetries and conserved quantities as demonstrated in the appendix of [78].

active field of research [82–85] and we may hope that an enhanced understanding of optical angular momentum might be helpful in the further development of the field.

4.1. Spin and orbital components of the angular momentum

The problem of separating optical angular momentum into spin and orbital parts has an interesting history. It has been suggested, on theoretical grounds, that such a separation is not physically meaningful [86–91], and such a conclusion is, at first sight, natural as these quantities are not separately conserved for the Dirac electron [48]. This position is somewhat at odds, however, with experimental work which suggested, at least in some circumstances, that precisely such a separation is meaningful [92, 93]. The resolution of this apparent contradiction is that the separation of the total angular momentum into spin and orbital parts is indeed a physical one, but that neither the spin nor the orbital components is, by itself, an angular momentum [94, 95]. Both the spin and orbital parts are the generators of rotations that are constrained by the requirement that the rotated fields must be transverse [38, 39].

We have seen that the total angular momentum density is expressed, in the formalism of the optical Dirac equation in terms of the alpha matrices in the form

$$\mathbf{j} = \psi^\dagger (\mathbf{r} \times \boldsymbol{\alpha}) \psi = \mathbf{r} \times (\mathbf{E} \times \mathbf{B}). \quad (72)$$

The spin and orbital parts of this total angular momentum are usually extracted by writing \mathbf{B} in terms of the vector potential [94, 95] or both \mathbf{B} in terms of the vector potential and \mathbf{E} in terms of a second vector potential [38, 39] and using integration by parts. Here we make this division directly from the properties of the optical Dirac equation.

At first sight it is far from obvious how to break up $\psi^\dagger (\mathbf{r} \times \boldsymbol{\alpha}) \psi$ into constant spin and orbital parts. We achieve this by making use of a property of the Kemmer matrices (A.11)

$$\begin{aligned} (\boldsymbol{\alpha} \cdot \mathbf{p})^2 &= p^2 - \mathcal{P}^2 \\ \Rightarrow \frac{(\boldsymbol{\alpha} \cdot \mathbf{p})^2}{p^2} &= \mathbf{I} - \frac{\mathcal{P}^2}{p^2}. \end{aligned} \quad (73)$$

We recall that $\mathcal{P}^2 \psi = 0$, by virtue of the transverse nature of the \mathbf{E} and \mathbf{B} fields, and hence for any physically allowed fields we have

$$\frac{(\boldsymbol{\alpha} \cdot \mathbf{p})^2}{p^2} \psi = \psi, \quad \psi^\dagger \frac{(\boldsymbol{\alpha} \cdot \mathbf{p})^2}{p^2} = \psi^\dagger. \quad (74)$$

It follows that we are free to rewrite the total angular momentum density in the form

$$\begin{aligned} \psi^\dagger (\mathbf{r} \times \boldsymbol{\alpha}) \psi &= \psi^\dagger (\mathbf{r} \times \boldsymbol{\alpha}) \frac{(\boldsymbol{\alpha} \cdot \mathbf{p})^2}{p^2} \psi \\ &= \psi^\dagger (\mathbf{r} \times \boldsymbol{\alpha}) \boldsymbol{\alpha} \cdot \mathbf{p} \frac{\boldsymbol{\alpha} \cdot \mathbf{p}}{p^2} \psi. \end{aligned} \quad (75)$$

We proceed by evaluating the effect on ψ of the operator $\boldsymbol{\alpha} \cdot \mathbf{p}/p^2$. In order to do so we introduce the transverse (divergenceless) and, *gauge-invariant* parts of two vector potentials, \mathbf{A}^\perp and \mathbf{C}^\perp , related to our (transverse) electric and magnetic fields by [9, 96]

$$\begin{aligned}\mathbf{E} &= -\frac{\partial}{\partial t}\mathbf{A}^\perp = -\nabla \times \mathbf{C}^\perp \\ \mathbf{B} &= -\frac{\partial}{\partial t}\mathbf{C}^\perp = \nabla \times \mathbf{A}^\perp.\end{aligned}\tag{76}$$

Hence with the aid of elementary vector identities we find

$$\phi = \frac{\boldsymbol{\alpha} \cdot \mathbf{p}}{p^2}\psi = \frac{1}{\hbar\sqrt{2}}\begin{pmatrix} i\mathbf{A}^\perp \\ -\mathbf{C}^\perp \end{pmatrix}.\tag{77}$$

We can complete our calculation of the total angular momentum density by acting with $(\mathbf{r} \times \boldsymbol{\alpha})\boldsymbol{\alpha} \cdot \mathbf{p}$ to the, *left* on ψ^\dagger . The result is

$$\psi^\dagger(\mathbf{r} \times \boldsymbol{\alpha})\psi = \psi^\dagger(\mathbf{r} \times \mathbf{p})\phi + \psi^\dagger\boldsymbol{\Sigma}\phi.\tag{78}$$

It is natural to identify these with the orbital, \mathbf{l} , and spin, \mathbf{s} parts of the angular momentum density and if we write these in terms of the fields and the potentials we find

$$\begin{aligned}\mathbf{l} &= \psi^\dagger(\mathbf{r} \times \mathbf{p})\phi = \frac{1}{2}\left(E_i\mathbf{r} \times \nabla A_i^\perp + B_i\mathbf{r} \times \nabla C_i^\perp\right) \\ \mathbf{s} &= \psi^\dagger\boldsymbol{\Sigma}\phi = \frac{1}{2}\left(\mathbf{E} \times \mathbf{A}^\perp + \mathbf{B} \times \mathbf{C}^\perp\right),\end{aligned}\tag{79}$$

which we recognize as the required Heaviside–Larmor symmetric forms [38, 39, 78]. Had we chosen to act with $(\mathbf{r} \times \boldsymbol{\alpha})\boldsymbol{\alpha} \cdot \mathbf{p}$ to the right we would have recovered $\mathbf{r} \times (\mathbf{E} \times \mathbf{B})$ and it follows that this expression and $\mathbf{l} + \mathbf{s}$, given in (79), are, *identical* within the formalism of the optical Dirac equation. We note that, as for the electron, this separation into spin and orbital parts is not Lorentz-covariant.

We do not wish to give the impression that the spin and orbital angular momentum densities given in (79) are unique. Had we written

$$\psi^\dagger(\mathbf{r} \times \boldsymbol{\alpha})\psi = \psi^\dagger \frac{(\boldsymbol{\alpha} \cdot \mathbf{p})^2}{p^2}(\mathbf{r} \times \boldsymbol{\alpha})\psi,$$

for example, then we would have obtained expressions for the densities differing only by a surface term, with the total volume-integrated spin and orbital angular momenta unchanged. The possibility of defining different densities for locally conserved quantities associated with the same total properties is familiar, of course, from Noether's theorem [76–78].

4.2. Helicity

In particle physics, the helicity of a single particle is defined as the component of its spin along its momentum [52, 97–100], thus it is natural to define this as the expectation value of $(\boldsymbol{\Sigma} \cdot \mathbf{p})/p$. If we apply this idea to our optical Dirac equation then we are led to a helicity density in the form

$$h_1 = \psi^\dagger \frac{\boldsymbol{\Sigma} \cdot \mathbf{p}}{p}\psi,\tag{80}$$

It is certainly, *possible* to define the helicity in this way and, indeed, doing so has some practical advantages [101]. There are two clear reasons, however, why this does not suit our present purposes. The first is that it gives a value that is proportional to \hbar which should not appear in

any physical quantity in our essentially *classical* analysis. The second, more serious, problem is that this quantity does not have the dimensions of an angular momentum density. We can remove both these features if we take our inspiration from (78) to give

$$h_2 = \psi^\dagger \frac{\boldsymbol{\Sigma} \cdot \mathbf{p}}{p} \psi, \quad (81)$$

but this quantity is a purely imaginary one.

We can recover the required form of the helicity density within from optical Dirac equation by defining the helicity operator to be $(\boldsymbol{\Sigma} \cdot \mathbf{p})/p^2$, so that the helicity density is

$$\begin{aligned} h &= \psi^\dagger \frac{\boldsymbol{\Sigma} \cdot \mathbf{p}}{p^2} \psi \\ &= \frac{1}{2} (\mathbf{A}^\perp \cdot \mathbf{B} - \mathbf{C}^\perp \cdot \mathbf{E}), \end{aligned} \quad (82)$$

which is the familiar expression [40–45, 78]. The total helicity obtained by integrating this helicity density over all space is the conserved quantity associated with the Heaviside–Larmor symmetry and, in the quantum theory, is \hbar multiplied by the difference between the number of left and right circularly polarized photons.

The difference in form for the helicity operator in the optical Dirac when compared with the analogous quantity for the electron should not come as a surprise, as this has been true for each of the mechanical properties we have considered and has its origins in the fact that $\psi^\dagger \psi$ has the dimensions of a probability density for the electron but an energy density for the electromagnetic field [16].

5. Zilch and higher-order conserved quantities

As a final application of the optical Dirac equation, we show how a class of further locally-conserved quantities, and in particular the Zilch [102, 103], arise naturally from it. These quantities have recently come to prominence in the study of chiral interactions [104–110].

We start by noting that if ψ is a solution of the optical Dirac equation, then so too is

$$\begin{aligned} \psi^{(1)} &= \frac{\boldsymbol{\Sigma} \cdot \mathbf{p}}{\hbar^2} \psi \\ &= \frac{1}{\sqrt{2}} \begin{pmatrix} \nabla \times \mathbf{E} \\ i \nabla \times \mathbf{B} \end{pmatrix}. \end{aligned} \quad (83)$$

The proof of this is straightforward, but instructive

$$\begin{aligned} i\hbar \frac{\partial}{\partial t} \psi^{(1)} &= \frac{\boldsymbol{\Sigma} \cdot \mathbf{p}}{\hbar^2} \boldsymbol{\alpha} \cdot \mathbf{p} \psi \\ &= \frac{(\boldsymbol{\alpha} \cdot \mathbf{p})^2}{\hbar^2 p^2} \boldsymbol{\Sigma} \cdot \mathbf{p} \boldsymbol{\alpha} \cdot \mathbf{p} \psi \\ &= \frac{(\boldsymbol{\alpha} \cdot \mathbf{p})^2}{\hbar^2 p^2} \begin{pmatrix} 0 & \mathbf{I} \\ \mathbf{I} & 0 \end{pmatrix} (\boldsymbol{\alpha} \cdot \mathbf{p})^2 \psi \end{aligned}$$

$$\begin{aligned}
&= \boldsymbol{\alpha} \cdot \mathbf{p} \frac{\boldsymbol{\Sigma} \cdot \mathbf{p}}{\hbar^2} \frac{(\boldsymbol{\alpha} \cdot \mathbf{p})^2}{p^2} \psi \\
&= \boldsymbol{\alpha} \cdot \mathbf{p} \psi^{(1)},
\end{aligned} \tag{84}$$

where we have made use of (A.5). It follows that we can derive locally conserved quantities using $\psi^{(1)}$ in place of ψ or, indeed, in combination with it. We present just one example

$$\psi^\dagger \psi^{(1)} = \frac{1}{2} [\mathbf{E} \cdot (\nabla \times \mathbf{E}) + \mathbf{B} \cdot (\nabla \times \mathbf{B})] = \frac{1}{2} Z^{000}, \tag{85}$$

where Z^{000} is one of the components of Lipkin's Zilch [102].

We can extend the family of analogous conserved quantities indefinitely [40, 44, 78] by noting that the spinors

$$\psi^{(n)} = \left(\frac{\boldsymbol{\Sigma} \cdot \mathbf{p}}{\hbar^2} \right)^n \psi \tag{86}$$

are also solutions of the optical Dirac equation. The repeated action of $\boldsymbol{\Sigma} \cdot \mathbf{p}/\hbar^2$ generates successive curl operations on the fields and thereby produces locally conserved quantities that depend on derivatives of the fields of ever higher order.

It is also possible to extend the set of conserved quantities in the other direction, with successive 'inverse-curl' operations to produce the infra-zilches [44, 78]. These arise naturally on the introduction of a further class of spinors

$$\psi^{(-n)} = \left(\frac{\boldsymbol{\Sigma} \cdot \mathbf{p}}{p^2} \right)^n \psi. \tag{87}$$

These are clearly solutions of the optical Dirac equation and denoted with the superscript $(-n)$ by virtue of the fact that

$$\frac{(\boldsymbol{\Sigma} \cdot \mathbf{p})^2}{\hbar^2 p^2} \psi = \psi, \tag{88}$$

so that the operators $\boldsymbol{\Sigma} \cdot \mathbf{p}/p^2$ acts on the $\psi^{(n)}$ as the, *inverse* of $\boldsymbol{\Sigma} \cdot \mathbf{p}/\hbar^2$. The simplest new spinor, $\psi^{(-1)}$, has the form

$$\psi^{(-1)} = \frac{\boldsymbol{\Sigma} \cdot \mathbf{p}}{p^2} \psi = \frac{1}{\sqrt{2}} \begin{pmatrix} -\mathbf{C}^\perp \\ i\mathbf{A}^\perp \end{pmatrix}. \tag{89}$$

The associated locally-conserved quantity $\psi^\dagger \psi^{(-1)}$ is simply the helicity density (82).

6. Conclusion

The formal connection between Maxwell's equations and the Dirac equation was noticed very soon after the latter was first derived and there is an extensive literature on the topic, much of it aimed at formulating a wavefunction for the photon [16, 21]. Our aim in this study has not been to address the quantum properties of light but rather to see what can be learnt about the classical electromagnetic field from the Dirac theory of the electron.

We have seen that many of the properties of light emerge naturally from the optical Dirac equation, including its principal mechanical properties and the associated conservation laws.

The familiar operation of cycle averaging optical fields, moreover, is precisely analogous to the suppression of electron, *Zitterbewegung* and that the introduction of complex fields arises naturally from the Foldy–Wouthuysen diagonalization of the optical Dirac ‘Hamiltonian’ $\boldsymbol{\alpha} \cdot \mathbf{p}$. The separation of the angular momentum into spin and orbital parts appears naturally in this formalism as do the infinite sets of conserved quantities: the zilches and infra-zilches.

Analogies between optics and quantum theory are well established, of course. Indeed Gabor once wrote that ‘optics was always considered as a good didactical preparation for wave mechanics, now it appears that quantum mechanics is not a bad preparation for optics’ [111]. Perhaps, *relativistic* quantum theory is an even better one.

Acknowledgements

I thank Les Allen, Rob Cameron, Sarah Croke, Fiona Speirits and Alison Yao for their constructive comments and encouragement. This work was supported by the Engineering and Physical Sciences Research Council (EPSRC) under grant number EP/I012451/1, by the Royal Society and by the Wolfson Foundation.

Appendix A. Properties of the Kemmer matrices

Useful properties of products of our spin-1 Kemmer (39) and (40) matrices follow from their commutators and anticommutators, the simplest of which is

$$\{\alpha_i, \beta\} = \alpha_i \beta + \beta \alpha_i = 0. \quad (\text{A.1})$$

The commutation relation for the alpha matrices is

$$\begin{aligned} [\alpha_i, \alpha_j] &= \frac{i}{\hbar} \varepsilon_{ijk} \Sigma_k \\ \Rightarrow [\mathbf{a} \cdot \boldsymbol{\alpha}, \mathbf{b} \cdot \boldsymbol{\alpha}] &= \frac{i}{\hbar} (\mathbf{a} \times \mathbf{b}) \cdot \boldsymbol{\Sigma}, \end{aligned} \quad (\text{A.2})$$

where ε_{ijk} is the familiar alternating symbol and we have employed the summation convention in which indices appearing twice are summed over the three cartesian directions. The matrix Σ_k is the k -component of the spin operator

$$\boldsymbol{\Sigma} = \hbar \begin{pmatrix} \boldsymbol{\tau} & 0 \\ 0 & \boldsymbol{\tau} \end{pmatrix}. \quad (\text{A.3})$$

Note the similarity with the electron operator (23). An important difference, of course, is that the electron operator has eigenvalues $\pm \hbar/2$ but that our optical spin operator has eigenvalues $\pm \hbar, 0$. These spin matrices commute with β

$$[\beta, \boldsymbol{\Sigma}] = 0. \quad (\text{A.4})$$

The operators $\boldsymbol{\alpha}$ and $\boldsymbol{\Sigma}$ are simply related by

$$\frac{1}{\hbar} \boldsymbol{\Sigma} = \begin{pmatrix} 0 & \mathbf{I} \\ \mathbf{I} & 0 \end{pmatrix} \boldsymbol{\alpha} = \boldsymbol{\alpha} \begin{pmatrix} 0 & \mathbf{I} \\ \mathbf{I} & 0 \end{pmatrix}. \quad (\text{A.5})$$

The anticommutator of two alpha matrices is of a more complicated form than the commutator, and follows from the anticommutator of the tau matrices

$$\left\{ \mathbf{a} \cdot \boldsymbol{\alpha}, \mathbf{b} \cdot \boldsymbol{\alpha} \right\} = \begin{pmatrix} \{\mathbf{a} \cdot \boldsymbol{\tau}, \mathbf{b} \cdot \boldsymbol{\tau}\} & 0 \\ 0 & \{\mathbf{a} \cdot \boldsymbol{\tau}, \mathbf{b} \cdot \boldsymbol{\tau}\} \end{pmatrix}, \quad (\text{A.6})$$

where $\{\mathbf{a} \cdot \boldsymbol{\tau}, \mathbf{b} \cdot \boldsymbol{\tau}\}$ is a symmetric three by three matrix with elements

$$\left\{ \mathbf{a} \cdot \boldsymbol{\tau}, \mathbf{b} \cdot \boldsymbol{\tau} \right\}_{ij} = 2\mathbf{a} \cdot \mathbf{b} \delta_{ij} - a_i b_j - b_i a_j. \quad (\text{A.7})$$

We can combine this with the commutation relation (A.2) to write the product of two τ matrices in the form

$$(\mathbf{a} \cdot \boldsymbol{\tau} \mathbf{b} \cdot \boldsymbol{\tau})_{ij} = \mathbf{a} \cdot \mathbf{b} \delta_{ij} - b_i a_j \quad (\text{A.8})$$

so that

$$\mathbf{a} \cdot \boldsymbol{\alpha} \mathbf{b} \cdot \boldsymbol{\alpha} = \begin{pmatrix} \mathbf{a} \cdot \boldsymbol{\tau} \mathbf{b} \cdot \boldsymbol{\tau} & 0 \\ 0 & \mathbf{a} \cdot \boldsymbol{\tau} \mathbf{b} \cdot \boldsymbol{\tau} \end{pmatrix}. \quad (\text{A.9})$$

It is also useful to note a simple expression for products of three matrices [59]

$$\alpha_i \alpha_j \alpha_k + \alpha_k \alpha_j \alpha_i = \delta_{jk} \alpha_i + \delta_{ij} \alpha_k. \quad (\text{A.10})$$

We conclude with a list of properties of the matrices in combination with the momentum

$$\begin{aligned} (\boldsymbol{\alpha} \cdot \mathbf{p})^2 &= p^2 - \mathcal{P}^2 \\ \frac{(\boldsymbol{\Sigma} \cdot \mathbf{p})^2}{\hbar^2} &= p^2 - \mathcal{P}^2 \\ (\boldsymbol{\alpha} \times \mathbf{p})^2 &= p^2 + \mathcal{P}^2 \\ \frac{(\boldsymbol{\Sigma} \times \mathbf{p})^2}{\hbar^2} &= p^2 + \mathcal{P}^2, \end{aligned} \quad (\text{A.11})$$

where \mathcal{P}^2 is the six by six matrix with dimensions p^2

$$\mathcal{P}^2 = \begin{pmatrix} \Pi & 0 \\ 0 & \Pi \end{pmatrix} \quad (\text{A.12})$$

and Π is the three by three symmetric matrix with elements

$$\Pi_{ij} = p_i p_j. \quad (\text{A.13})$$

Our ψ contains only transverse fields and this means that \mathcal{P}^2 acting on ψ give zero

$$\mathcal{P}^2 \psi = 0, \quad \psi^\dagger \mathcal{P}^2 = 0. \quad (\text{A.14})$$

Appendix B. Optical Foldy–Wouthuysen transformation of the Hamiltonian

We seek a unitary transformation that diagonalizes the Hamiltonian operator for our optical Dirac equation. The transformed Hamiltonian has the form

$$\begin{aligned}
H' &= \exp(\theta\beta\boldsymbol{\alpha} \cdot \mathbf{p})\boldsymbol{\alpha} \cdot \mathbf{p} \exp(-\theta\beta\boldsymbol{\alpha} \cdot \mathbf{p}) \\
&= \exp(\theta\beta\boldsymbol{\alpha} \cdot \mathbf{p})\beta^2\boldsymbol{\alpha} \cdot \mathbf{p} \exp(-\theta\beta\boldsymbol{\alpha} \cdot \mathbf{p}) \\
&= -\exp(\theta\beta\boldsymbol{\alpha} \cdot \mathbf{p})\beta\boldsymbol{\alpha} \cdot \mathbf{p}\beta \exp(-\theta\beta\boldsymbol{\alpha} \cdot \mathbf{p}) \\
&= -\beta\boldsymbol{\alpha} \cdot \mathbf{p} \exp(\theta\beta\boldsymbol{\alpha} \cdot \mathbf{p})\beta \exp(-\theta\beta\boldsymbol{\alpha} \cdot \mathbf{p}) \\
&= -\beta\boldsymbol{\alpha} \cdot \mathbf{p}\beta \exp(-2\theta\beta\boldsymbol{\alpha} \cdot \mathbf{p}) \\
&= \boldsymbol{\alpha} \cdot \mathbf{p} \exp(-2\theta\beta\boldsymbol{\alpha} \cdot \mathbf{p}).
\end{aligned} \tag{B.1}$$

We can use the expanded form of the unitary operator (51) to write this as

$$H' = \boldsymbol{\alpha} \cdot \mathbf{p} \cos(2p\theta) + \beta \frac{(\boldsymbol{\alpha} \cdot \mathbf{p})^2}{p} \sin(2p\theta). \tag{B.2}$$

To get a diagonal form we need to remove the $\boldsymbol{\alpha} \cdot \mathbf{p}$ part and to this end we choose

$$\theta = \frac{\pi}{4p}. \tag{B.3}$$

The desired Hamiltonian is then

$$H' = \beta \frac{(\boldsymbol{\alpha} \cdot \mathbf{p})^2}{p} = \beta \left(p - \frac{p^2}{p} \right). \tag{B.4}$$

References

- [1] Maxwell J C 1865 A dynamical theory of the electromagnetic field *Phil. Trans. R. Soc.* **155** 459–512
- [2] Heaviside O 1971 *Electromagnetic Theory* (Providence, RI: AMS Chelsea Publishing)
- [3] Landau L D and Lifshitz E M 1975 *The Classical Theory of Fields* (Oxford: Butterworth-Heinemann)
- [4] Weinberg S 1972 *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity* (New York: Wiley)
- [5] Weber H 1901 *Die Partiellen Differential-Gleichung der Mathematischen Physik nach Riemann's Vorlesungen* vol 2 (Braunschweig: Friedrich Vieweg und Sohn)
- [6] Silberstein L 1907 Elektromagnetische Grundgleichungen in bivectorieller Behandlung *Ann. Phys., Lpz.* **22** 579–86
- [7] Silberstein L 1907 Nachtrag zur Abhandlung über 'elektromagnetische Grundgleichungen in bivectorieller Behandlung' *Ann. Phys., Lpz.* **24** 783–4
- [8] Silberstein L 1914 *The Theory of Relativity* (London: Macmillan)
- [9] Bateman H 1915 *The Mathematical Analysis of Electrical and Optical Wave-Motion* (Cambridge: Cambridge University Press) Reprinted 1955 (New York: Dover)
- [10] Stratton J A 1941 *Electromagnetic Theory* (New York: McGraw-Hill)
- [11] Power E A and Thirunamachandran T 1999 Two derivations of the Feynman formulae for the fields of a moving charge *J. Mod. Opt.* **46** 1869–79
- [12] Oppenheimer J R 1931 Note on light quanta and the electromagnetic field *Phys. Rev.* **38** 725–46
- [13] Molière G 1949 Laufende elektromagnetische Multipolwellen und eine neue methods der Feld-Quantisierung *Ann. Phys., Lpz.* **6** 146–62
- [14] Archibald W J 1955 Field equations from particle equations *Can. J. Phys.* **33** 565–72
- [15] Białynicki-Birula I 1994 On the wave function of the photon *Acta Phys. Pol.* **86** 97–116
- [16] Białynicki-Birula I 1996 Photon wave function *Progress in Optics* 36th edn ed E Wolf (Amsterdam: Elsevier) pp 245–94

- [17] Schiff L I 1968 *Quantum Mechanics* (Tokyo: McGraw-Hill)
- [18] Weyl H 1929 Elektron und gravitation I *Z. Phys.* **56** 330–52
- [19] Good R H 1957 Particle aspect of the electromagnetic field equations *Phys. Rev.* **105** 1914–9
- [20] Sipe J E 1995 Photon wave functions *Phys. Rev. A* **52** 1875–83
- [21] Smith B J and Raymer M G 2007 Photon wave functions, wave-packet quantization of light and coherence theory *New J. Phys.* **9** 414
- [22] Landau L and Peierls R 1930 Quantenelektrodynamik im Konfigurationsraum *Z. Phys.* **62** 188–200
- [23] Cook R J 1982 Photon dynamics *Phys. Rev. A* **25** 2164–7
- [24] Cook R J 1982 Lorentz covariance of photon dynamics *Phys. Rev. A* **26** 2754–60
- [25] Inagaki T 1994 Quantum-mechanical approach to a free photon *Phys. Rev. A* **49** 2839–43
- [26] Scully M O and Zubairy M S 1997 *Quantum Optics* (Cambridge: Cambridge University Press)
- [27] Darwin C G 1932 Notes on the theory of radiation *Proc. R. Soc. A* **136** 36–52
- [28] Mohr P J 2010 Solutions of the Maxwell equations and photon wave functions *Ann. Phys., NY* **325** 607–63
- [29] Caran E 1966 *The Theory of Spinors* (Paris: Hermann) Reprinted 1981 (New York: Dover)
- [30] Corson E M 1953 *Introduction to Tensors, Spinors, and Relativistic Wave Equations* (London: Blackie and Son)
- [31] Penrose R and Rindler W 1984 *Spinors and Space-Time* vol 1 (Cambridge: Cambridge University Press)
- [32] Laporte O and Uhlenbeck G E 1931 Application of spinor analysis to the Maxwell and Dirac equations *Phys. Rev.* **37** 1380–97
- [33] Moses H E 1959 Solution of Maxwell's equations in terms of a spinor notation: the direct and inverse problem *Phys. Rev.* **113** 1670–9
- [34] Heaviside O 1892 On the forces, stresses and fluxes of energy in the electromagnetic field *Phil. Trans. R. Soc. A* **183** 423–80
- [35] Larmor J 1897 Dynamical theory of the electric and lumiferous medium III *Phil. Trans. R. Soc. A* **190** 205–300
- [36] Jackson J D 1999 *Classical Electrodynamics* 3rd edn (New York: Wiley)
- [37] Rose M E 1955 *Multipole Fields* (New York: Wiley)
- [38] Barnett S M 2010 Rotation of electromagnetic fields and the nature of optical angular momentum *J. Mod. Opt.* **57** 1339–43
- [39] Barnett S M 2011 On the six components of optical angular momentum *J. Opt.* **13** 064010
- [40] Candlin D J 1965 Analysis of the new conservation law in electromagnetic theory *Il Nuovo Cimento* **37** 1390–5
- [41] Calkin M G 1965 An invariance property of the free electromagnetic field *Am. J. Phys.* **33** 958–60
- [42] Zwanziger D 1968 Quantum field theory of particles with both electric and magnetic charge *Phys. Rev.* **176** 1489–95
- [43] Afanasiev G N and Stepanovsky Y P 1996 The helicity of the free electromagnetic field and its physical meaning *Il Nuovo Cimento* **109** 271–9
- [44] Cameron R P, Barnett S M and Yao A M 2012 Optical helicity, optical spin and related quantities in electromagnetic theory *New J. Phys.* **14** 053050
- [45] Barnett S M, Cameron R P and Yao A M 2012 Duplex symmetry and its relation to the conservation of optical helicity *Phys. Rev. A* **86** 013845
- [46] Barnett S M 2014 Maxwellian theory of gravitational waves and their mechanical properties *New J. Phys.* **16** 023027
- [47] Høegåsen H 1962 A study in Kemmer-algebra and Foldy–Wouthuysen transformations for particles with spin 0 and 1 *Proc. R. Norwegian Acad. Sci.* **6** 1–53
- [48] Dirac P A M 1928 The quantum theory of the electron *Proc. R. Soc. A* **117** 610–24
- [49] Dirac P A M 1958 *The Principles of Quantum Mechanics* (Oxford: Oxford University Press)
- [50] Rose M E 1961 *Relativistic Electron Theory* (New York: Wiley)
- [51] Bjorken J D and Drell S D 1964 *Relativistic Quantum Mechanics* (New York: McGraw-Hill)

- [52] Aitchison I J R 1972 *Relativistic Quantum Mechanics* (London: Macmillan)
- [53] Itzykson C and Zuber J-B 1980 *Quantum Field Theory* (London: Macmillan)
- [54] Bethe H A and Salpeter E E 1957 *Quantum Mechanics of One-and Two-Electron Atoms* (New York: Academic)
- [55] Foldy L L and Wouthuysen S A 1950 On the Dirac theory of spin 1/2 particles and its non-relativistic limit *Phys. Rev.* **78** 29–36
- [56] Pryce M H L 1935 Commuting co-ordinates in the new field theory *Proc. R. Soc. A* **150** 166–72
- [57] Pryce M H L 1948 The mass-centre in the restricted theory of relativity and its connexion with the quantum theory of elementary particles *Proc. R. Soc. A* **195** 62–81
- [58] Majorana E 1932 Teoria relativistica di particelle con momento intrinseco arbitrario *Nuovo Cimento* **9** 335–44
- [59] Kemmer N 1939 The particle aspect of meson theory *Proc. R. Soc. A* **173** 91–116
- [60] Duffin R J 1938 On the characteristic matrices of covariant systems *Phys. Rev.* **54** 1114
- [61] Kemmer N 1943 The algebra of meson matrices *Proc. Camb. Phil. Soc* **39** 189–96
- [62] Schrödinger E 1943 Pentads, tetrads, and traids of meson-matrices *Proc. R. Ir. Acad.* **48** 135–46
- [63] Schrödinger E 1943 Systematics of meson-matrices *Proc. R. Ir. Acad.* **49** 29–42
- [64] Schrödinger E 1955 The wave equation for spin 1 in Hamiltonian form *Proc. R. Soc. A* **229** 39–43
- [65] Schrödinger E 1955 The wave equation for spin 1 in Hamiltonian form II *Proc. R. Soc. A* **232** 435–47
- [66] Landau L D 1941 The theory of superfluidity of Helium II *J. Phys* **5** 71–90 Reprinted in [67]
- [67] Khalatnikov I M 2000 *An Introduction to the Theory of Superfluidity* (Perseus: Cambridge MA)
- [68] Whitlock N K, Barnett S M and Jeffers J 2004 On the mechanical properties of matter waves *J. Phys. B: At. Mol. Opt. Phys.* **37** L293–6
- [69] Barnett S M 2002 Optical angular-momentum flux *J. Opt. B: Quantum Semiclass. Opt.* **4** S7–16
- [70] Glauber R J 1963 The quantum theory of optical coherence *Phys. Rev.* **130** 2529–39
- [71] Glauber R J 2007 *Quantum Theory of Optical Coherence* (New York: Wiley)
- [72] Loudon R 2000 *The Quantum Theory of Light* 3rd edn (Oxford: Oxford University Press)
- [73] Siegman A E 1986 *Lasers* (Sausalito CA: University Science Books)
- [74] Milonni P W and Eberly J H 2010 *Laser Physics* (Hoboken, NJ: Wiley)
- [75] Born M and Wolf E 1987 *Principles of Optics* 6th edn (Oxford: Pergamon)
- [76] Noether E 1918 Invariante variationsprobleme *Nachr. D. Königl. Gesellsch. D. Wiss. Zu Göttingen Math.-Phys. Kl.* **2** 235–57
- [77] Neuenschwander D E 2011 *Emmy Noether's Wonderful Theorem* (Baltimore, MD: John Hopkins University Press)
- [78] Cameron R P and Barnett S M 2012 Electric-magnetic symmetry and Noether's theorem *New J. Phys.* **14** 123019
- [79] Bliokh K Y, Bekshaev A Y and Nori F 2013 Dual electromagnetism: helicity, spin, momentum and angular momentum *New J. Phys.* **15** 033026
- [80] Bekshaev A, Bliokh K Y and Soskin M 2011 Internal flows and energy circulation in light beams *J. Opt.* **13** 053001
- [81] Bohm D 1951 *Quantum Theory* (Englewood Cliffs, NJ: Prentice-Hall) Reprinted 1989 (New York: Dover)
- [82] Allen L, Beijersbergen M W, Spreeuw R J C and Woerdman J P 1992 Orbital angular momentum of light and the transformation of Laguerre-Gaussian laser modes *Phys. Rev. A* **45** 8185–9
- [83] Allen L, Barnett S M and Padgett M J 2003 *Optical Angular Momentum* (Bristol: Institute of Physics Publishing)
- [84] Yao A M and Padgett M J 2010 Orbital angular momentum: origins, behaviour and applications *Adv. Opt. Photonics* **2** 161–204
- [85] Andrews D L and Babiker M 2013 *The Angular Momentum of Light* (Cambridge: Cambridge University Press)
- [86] Cohen-Tannoudji C, Dupont-Roc J and Grynberg G 1989 *Photons and Atoms* (New York: Wiley)

- [87] Jauch J M and Rohrlich F 1955 *The Theory of Photons and Electrons* (Cambridge, MA: Addison-Wesley)
- [88] Rohrlich F 1965 *Classical Charged Particles* (Reading, MA: Addison-Wesley)
- [89] Yilmaz H 1965 *Introduction to the Theory of Relativity and Principles of Modern Physics* (New York: Blaisdell)
- [90] Simmons J W and Guttman M J 1970 *States, Waves and Photons* (Reading, MA: Addison-Wesley)
- [91] Barut A O 1980 *Electrodynamics and Classical Theory of Fields and Particles* (New York: Dover)
- [92] He H, Friese M E J, Heckenberg N R and Rubinsztein-Dunlop H 1995 Direct observation of transfer of angular momentum to absorptive particles from a laser beam with a phase singularity *Phys. Rev. Lett.* **75** 826–9
- [93] Simpson N B, Dholakia K, Allen L and Padgett M J 1997 Mechanical equivalence of spin and orbital angular momentum of light: an optical spanner *Opt. Lett.* **22** 52–4
- [94] van Enk S J and Nienhuis G 1994 Commutation rules and eigenvalues of spin and orbital angular momentum of radiation fields *J. Mod. Opt.* **41** 963–77
- [95] van Enk S J and Nienhuis G 1994 Spin and orbital angular momentum of photons *Europhys. Lett.* **25** 497–501
- [96] Cameron R P 2014 On the ‘second potential’ in electrodynamics *J. Opt.* **16** 015708
- [97] Schweber S S 1961 *An Introduction to Relativistic Quantum Field Theory* (New York: Harper and Row) Reprinted 2005 (New York: Dover)
- [98] Aitchison I J R and Hey A J G 1989 *Gauge Theories in Particle Physics* 2nd edn (Bristol: Adam Hilger)
- [99] Leader E 2001 *Spin in Particle Physics* (Cambridge: Cambridge University Press)
- [100] Maggiore M 2005 *A Modern Introduction to Quantum Field Theory* (Oxford: Oxford University Press)
- [101] Białynicki-Birula I and Białynicka-Birula Z 2009 Why photons cannot be sharply localised *Phys. Rev. A* **79** 032112
- [102] Lipkin D M 1964 Existence of a new conservation law in electromagnetic theory *J. Math. Phys.* **5** 696–700
- [103] Kibble T W B 1965 Conservation laws for free fields *J. Math. Phys.* **6** 1022–6
- [104] Tang Y and Cohen A E 2010 Optical chirality and its interaction with matter *Phys. Rev. Lett.* **104** 163901
- [105] Hendry E, Carpy T, Johnston T, Popland M, Mikhaylovskiy R V, Lapthorn A J, Kelly S M, Barron L D, Gadegaard N and Kadodwala M 2010 Ultrasensitive detection and characterisation of biomolecules using superchiral fields *Nat. Nanotechnology* **5** 783–7
- [106] Tang Y and Cohen A E 2011 Enhanced enantioselectivity in excitation of chiral molecules by superchiral light *Science* **332** 333–6
- [107] Bliokh K Y and Nori F 2011 Characterizing optical chirality *Phys. Rev. A* **83** 021803
- [108] Coles M M and Andrews D L 2012 Chirality and angular momentum in radiation *Phys. Rev. A* **85** 063810
- [109] Andrews D L and Coles M M 2012 Measures of chirality and angular momentum in the electromagnetic field *Opt. Lett.* **37** 3009–11
- [110] Rosales-Guzmán C, Volke-Sepulveda K and Torres J P 2012 Light with enhanced optical chirality *Opt. Lett.* **37** 3486–8
- [111] Gabor D 1956 Light and information *Proceedings of a Symposium on Astronomical Optics and Related Subjects* ed Z Kopal (Amsterdam: North-Holland)